

Determining Company-Specific Legitimate Domestic Night Use Data

Paul Fanner, Technical Services Manager, Bristol Water and
Dr Charles Harris, Silkleaf Design

1. Introduction

The level of leakage calculated from DMA nightline data is sensitive to the estimated levels of household and non-household legitimate night use. Until recently, most water companies have relied on the Managing Leakage default figure of 1.7 litres/property/hour for estimating legitimate domestic night use. The use of national default figures is no longer acceptable because OFWAT will not accept a Company's own leakage targets unless they are based on company-specific data.

In order to effectively manage active leakage control work from DMA nightline data, DMA-specific domestic night use figures would ideally be obtained directly for each DMA. However, in practice it is not possible to determine this data directly from DMA data because the DMA nightlines will also include legitimate non-household night use. Bristol Water has therefore used data from small Consumption Monitor Areas (CMAs) to estimate company-specific domestic night use.

This paper describes how Socrates data from CMAs has been analysed to calculate site-specific customer night use figures expressed in litres/property/hour. It also summarises the results obtained to date.

The paper uses the example of data taken from one of Bristol Water's consumption monitor areas, Chessel Street CMA, which includes 956 properties.

2. The Consumption Monitors

The first phase of the Bristol Water consumption monitor consisted of 30 small consumption monitor areas, ranging in size from 70 to 987 properties. These CMAs were selected to be statistically representative of the Company's customer base, and were selected, as far as possible, to have a single ACORN classification, a minimum of non-household use and a minimum of metered domestic customers. These CMAs have been used in the determination of site-specific legitimate domestic night use data. Because of the statistical nature of these CMAs, it is possible to apply the results obtained from the CMAs to the whole Company's area. It is also possible to apply the results at DMA level using the ACORN classifications within the DMA. The level of legitimate household and non-household night use are important factors in the determination of PCC from the consumption monitors. It is also possible to apply the results at DMA level using the ACORN classifications within the DMA.

3. The Data

For the determination of domestic night use, Socrates nightline analysers are installed on the CMA meter flow and pressure tapplings, in parallel with the permanent consumption logger. For each night it is deployed, Socrates collects the following data from the analysis of a two-hour nightline window:

- i) a minimum flow value;
- ii) an average flow value calculated over the two hour period;
- iii) an average pressure value calculated over the two hour period.

The minimum values are selected to be as representative as possible of genuine minimum flows for each night undistorted by surge flows or measurement errors.

It is recommended that approximately 30 nights' data are analysed. The consumption monitor area should be fully isolated from neighbouring networks and there should be no maintenance intervention during the analysis period. For the purposes of calculating consumption, the monitor area should have a reasonably stable rate of leakage.

4. The Principle

4.1 Core Flow

A consumption monitor nightline comprises intermittent user 'events' (toilet flushes, washing machine fills, etc.) superimposed on a 'core' flow level which will be very close to the actual leakage level. With small consumption monitor areas (up to approximately 250 properties) Socrates will record core flow levels because it measures over short periods, and thus penetrates between the user events defined above. An example is shown in section 4.2, below.

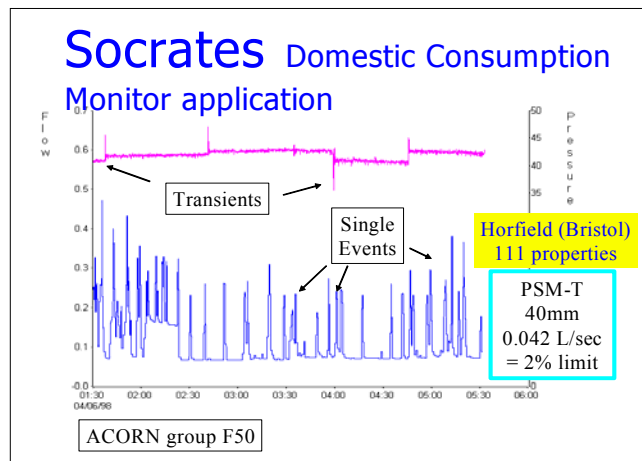


Figure 1 - Socrates Penetration to Core Flow

4.2 Why Does Core Flow Vary?

Figure 2 shows the distribution of core flows over a sequence of nightlines for the Chessel Street consumption monitor area. A nightly variation is apparent. What causes this variation?

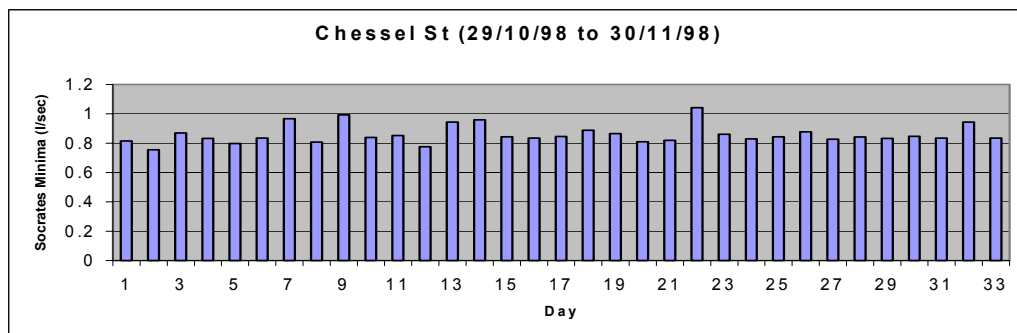


Figure 2 - Nightly Variation in Socrates Minima

Core flow comprises components of leakage flow, continuous 'use', and plumbing losses. The model proposes that leakage flow is relatively stable over the analysis period. Continuous use, however, varies randomly – an example being flow from a tap which has not been turned off properly; one night it may be left dripping rapidly, the next night it is fully off. Plumbing losses (due to cistern overflows, faulty tap washers etc) also happen randomly. The model proposes that it is these effects that cause the nightly variation in core flow. Figure 3 illustrates the components described above.

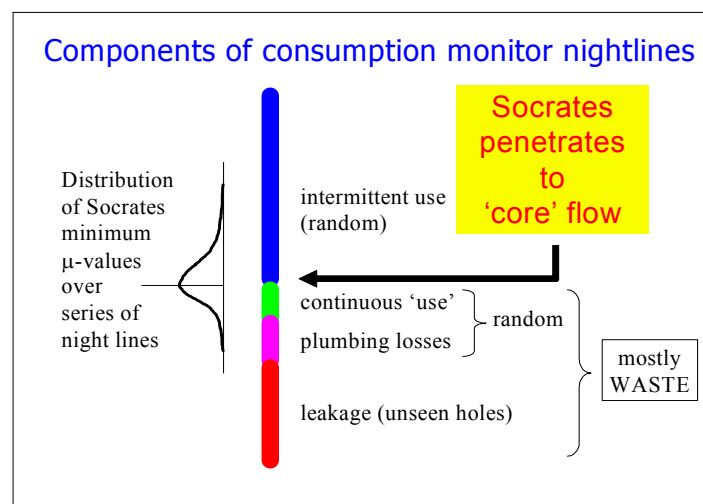


Figure 3 - Components of Core Flow

4.3 Pressure Effects

Pressure variation may also be a contributing factor but its effect has been found relatively minor. Table 1 shows instances of the unpredictable effect of pressure variation on core flow in which we see large changes in flow accompanied only by minor and wrongly phased changes in pressure. This implies that use is the dominant factor affecting the minimum flow rather than pressure variations. Despite the small effect, pressure variation is nevertheless taken into account in the analysis.

Date	Minimum Flow (l/sec)	Variation in Flow From Night to Night Relative to Mean Flow (%)	Pressure at Minimum Flow (m head)	Variation in Pressure Night to Night Relative to Mean Pressure (%)
14/8/98	0.091		47.67	
15/8/98	0.055	-49 %	47.78	+0.2 %
16/8/98	0.076	+29 %	47.69	-0.2 %
17/8/98	0.075	-1 %	48.13	+0.9 %
18/8/98	0.065	-14 %	46.79	-2.8 %
19/8/98	0.075	+14 %	46.93	+0.3 %
20/8/98	0.076	+1 %	47.89	+0.2 %

Table 1 – Nightly Variation in Minimum Flow and Pressure (Hotwells CMA)

To summarise: Core flow levels vary because of randomly changing use and loss components superimposed on the leakage level. A statistical estimate of the leakage level may be obtained by analysing the variation of the random components.

5 The Statistical Method

5.1 The Core Flow Distribution

Normally we never see the actual leakage level in a system because it is buried beneath random low level continuous use or losses. If, however, we were to observe the system for a very long time *and* the leakage state did not change over that time, sooner or later we would observe a measurement in which there was no continuous use or plumbing loss. This measurement would be the leakage level. Practically, we cannot observe systems over indefinite periods in which the leakage state is static. At best we have but 30 or so night's worth of data in which to observe the monitor area. But the distribution of that data indicates the likely position where random continuous use tails off to insignificant levels. That position becomes our best estimate of steady state leakage.

The Socrates minima - the core flow levels - will be distributed about some mean value μ . The spread of the distribution can be measured by its standard deviation σ . We can estimate the position of the likely leakage level as being some multiple of σ down from the mean. In our analysis we take the leakage position as being 2.5 σ down from the mean. The number 2.5 has been chosen from a WRc suggested range of 2.5 to 3. We use 2.5 as being the most cautious value in that range.

5.2 Skew in the Distributions

Distributions of Socrates minima are rarely symmetrical. Sometimes they are skewed positively, at other times negatively. There are two ways to mitigating the effects of this skew.

5.2.1 Outliers

Some nightlines appear abnormally high compared with their peers in the sequence of data being evaluated. Examples may be seen in Figure 2. For example, night 22 is significantly higher than most of the nights in the sequence. A reason for this may be an unusual amount of use during that night. If we include such outliers in our analysis, the mean level of the distribution rises but so too does its standard deviation. The net effect is to generally produce a leakage estimate ($= \mu - 2.5 \sigma$) which is unduly low because of the positive skew introduced by including the outlier. In producing an estimate of leakage, we need a representative set of minimum values that describe the usual pattern of events in the area. Unusually high values distort that estimate. WRc have suggested that positive outliers may be identified as being

those values which have a standardised residual between 2 and 3. A value's standardised residual is the value minus the mean all divided by the standard deviation. Again, erring on the cautious side, we reject values with outliers with residuals greater than 2.

Negative outliers are also possible. Some minimum values may appear unduly low. Causes include surge flows which have not been trapped by Socrates' filters. Low minimum values however may be due to genuine periods of low use. It is possible to inspect the lowest minimum values using Socrates' snapshot facility to ascertain whether these values are likely to be genuine or false. False values are manually removed.

5.2.2 Residual Skew

Even after outliers have been removed, it is rare for a distribution of minima to be symmetrical about its mean; a degree of positive or negative skew may persist. On some distributions it is possible to reduce the effect of skew by transforming the distribution to a log distribution. To do this, naperian logarithms are taken of each minimum value, and the mean and standard deviation calculated for these transformed values. The log of the estimate of the position of the leakage for the transformed distribution is the mean of the log values minus 2.5 times the standard deviation of the log values. The linear estimate (i.e. non log value) is then taken as the anti-log of the log estimate.

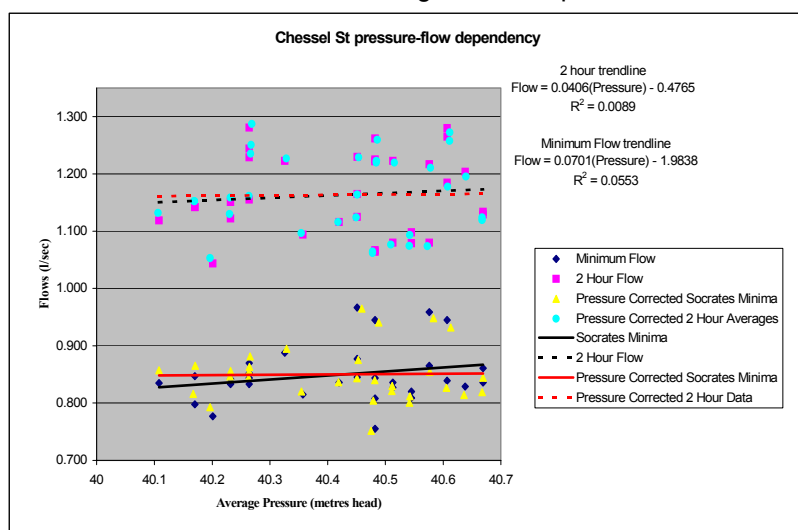
To decide whether to use a linear or log estimate, the system calculates the skew of both types of distribution using Pearson's second coefficient of skew. This coefficient is defined as three times the difference of the median minus the mean all divided by the standard deviation. Modes are often difficult to locate which is why the median is used in this calculation. The distribution with a skew value closest to zero is chosen.

6 Calculating Nightly Usage

Once the estimate of leakage level has been calculated for the sequence of nights, a value for each night's use can be calculated by subtracting the leakage level from the two hour average flow value. The two hour average flow value is used to be consistent with the Managing Leakage methodology for obtaining night use estimates. The average night use is obtained as the mean of all the nights' usage. Night use is expressed in litres/property/hour.

7 Compensating for Pressure Effects

Below is a scatter diagram showing each night's minimum and average flow plotted against corresponding average pressure. The lower cluster of points (diamonds and triangles) show the correlation of minimum flow with pressure. The upper set (squares and circles) show the correlation between two hour average flow and pressure. In this example, there is slight



tendency for both minimum and two hour average flows to rise with increasing pressure. This, however, is not always the case. Sometimes friction effects predominate and the reverse happens. Neither is the degree of correlation predictable from site to site, this being yet another instance where using a global average relationship may be misleading.

Figure 4 - Compensating for Pressure Effects

Because of this variability, the Socrates analysis characterises the relationship between flow and pressure for each set of measurements. Consider, for instance, the lower cluster of data points. The diamonds show the coordinates of the uncorrected average pressure and

minimum flow points. The sloping line through these points is a least squares best fit straight line. A straight line is justified here because we are considering small variations about an operating point. On the big 24-hour scale, the pressure-flow relationship may be highly non-linear, but here we are considering small perturbations about the mean nightline operating position. The slope of this line enables us to calculate how we must adjust our flow values to make them independent of pressure. The adjusted values are shown as triangles. The trend line through these transformed points is flat as intended.

A similar procedure can be carried out on the two hour flow points. The adjusted points here are shown as circles. Both sets of data are pressure adjusted about the mean site nightline operating pressure. Pressure adjusted values are used in the calculation described in section 5 above.

8 Corrections for Large Site Effects

Up to approximately 250 properties, Socrates always cuts through intermittent use to reveal the core flow level. As property numbers grow beyond this, individual user events start to merge together thus obscuring the core flow level. Simulations have been carried out to quantify this effect and to identify the likely 'penetration' error at each size of DMA.

Detailed four hour nightline flow measurements were taken over two week periods from six consumption monitor sites. Using data from each site, large DMAs were simulated by adding together randomly selected two hour nightline segments. Because the core flow level was visible in each data segment, the effect on combined core level visibility could be characterised as segments were merged together.

For each site, 10 random simulations were carried out testing 5 different averaging times per simulation for 8 property counts (500 to 4000). The results are shown in Table 2.

Property Count	Averaging Time (seconds)				
	15	30	60	90	120
500	0.016	0.017	0.020	0.023	0.025
1000	0.047	0.051	0.061	0.073	0.085
1500	0.098	0.107	0.129	0.146	0.164
2000	0.150	0.165	0.193	0.216	0.242
2500	0.217	0.236	0.269	0.300	0.327
3000	0.293	0.314	0.350	0.392	0.425
3500	0.356	0.384	0.427	0.474	0.517
4000	0.422	0.458	0.507	0.569	0.605

Table 2 - Average Differences Between Actual Minima and Socrates Minima for Different Property Counts and Averaging Times (litres/second)

A practical example illustrates the use of this table. Consider a Socrates minimum measurement taken on a 1000 property DMA using an average time of 30 seconds. The table suggests that the measured value will be too high by 0.051L/sec. This error is called the penetration error.

9 Results

Table 4 details the results of Socrates analysis of CMA data undertaken to date. It can be seen from the table that some CMAs have been analysed over 2 or 3 different periods to provide an analysis of seasonal variation in domestic night use. The night use data obtained from this work shows significant variability, both at the same site over time and between sites within the same ACORN band. These results confirm the need to obtain seasonal data from each site as there has been a consistent increase in night use seen during the summer period.

The results obtained indicate that household night use is considerably higher than the Managing Leakage default. If all the data obtained to date is used, the property weighted average of the effective minima analysis is 3.25 l/prop/hr and if the Socrates statistical analysis were used, legitimate night use calculated on a property weighted average would increase to 3.69 l/prop/hr.

Extrapolation of the effective minima analysis across the Company using ACORN categories results in an estimate of legitimate household night use of 3.44 l/prop/hr, if all the data is used. In practice, because we still only have a limited data set, Bristol Water is taking a cautious approach to using this data. We are currently eliminating outlying high values, such as those recorded at Winterbourne and Southville, until more data is available to confirm these values.

10 Conclusions

This project has demonstrated that nightflow data from small area household consumption monitors can be used to estimate legitimate household night use if Socrates is used to differentiate between leakage and legitimate use. The data collected exhibits significant variability, with a consistent increase in summer night use. However the data clearly indicates that the Managing Leakage default under-estimates actual household night use.

ACORN Type	CMA No.	Site Name	No. of Props	Dates of Data	No. of Days	l/prop/hr Using Minima	l/prop/hr Effective Minima	l/prop/hr Using Statistics
A01	4045	Tickenham	218	19.5 - 25.6	38	2.52	4.45	5.79
A03	7022	Bishop Sutton	284	18.5 - 28.6	42	1.77	4.46	7.75
B10	1094	Wickwar	188	5.5 - 28.6	55	2.11	2.57	2.77
B13	2131	Little Stoke	275	8.1 - 31.1	24	2.36	2.99	3.13
B13	2131	Little Stoke2	275	1.7 - 18.8	49	3.11	4.97	5.16
B14	1133	Frampton	268	29.10 - 7.12	40	2.15	2.5	2.74
B14	1138	Winterbourne	188	9.7 - 18.8	41	4.61	6.96	7.84
B15	8310	Weston	89	18.12 - 21.1	35	1.54	1.75	1.82
B15	8310	Weston2	89	6.2 - 10.3	33	1.47	1.67	1.99
C21	8311	Hotwells	132	17.11 - 17.12	31	1.14	1.44	1.55
C21	8311	Hotwells2	132	13.1 - 10.2	29	1.28	1.75	2.12
D27	8313	Clutton2	114	6.1 - 18.2	44	0.33	0.72	1.37
D30	8316	Downend	96	19.12 - 11.1	24	1.3	1.75	2.14
D30	8316	Downend2	96	12.1 - 4.3	52	1.21	1.57	1.71
D31	2246	Henbury	163	19.12 - 19.1	32	3.77	4.22	4.43
D31	2246	Henbury2	163	26.1 - 10.3	44	2.82	3.52	3.78
D31	2246	Henbury3	163	28.5 - 30.6	34	4.2	4.9	5.31
D31	3172	Chessel2	959	9.1 - 23.2	46	1.15	1.62	1.85
D31	2268	Ashley Down	212	21.4 - 21.6	62	2.59	3.83	4.2
D26	2132	Patchway	778	15.5 - 30.6	47	1.88	2.27	3.15
E33	3044	Stockwood	987	04.03 - 07.04	35	1.07	2.08	1.77
E33	3044	Stockwood2	987	17.7 - 18.8	33	1.28	2.2	2.41
E33	8319	Hartcliffe	111	11.11 - 15.12	35	2.77	4.13	4.97
E33	8319	Hartcliffe2	111	16.12 - 9.2	56	2.05	2.73	4.36
E33	8319	Hartcliffe3	111	10.2 - 10.3	29	1.06	1.4	1.52
E34	2275	Eastville	482	13.2 - 11.3	27	2.35	3.33	3.36
E34	2275	Eastville2	482	22.6 - 9.8	25	2.48	3.54	4.39
E37	3112	Totterdwon	670	27.7 - 20.8	25	2.4	3.23	3.75
E37	3174	Southville	877	11.03 - 12.04	33	5.84	7.25	8.07
E38	2277	St Werburghs	416	6.5 - 21.6	47	1.98	4.21	4.26
F40	3416	Lawrence	616	15.7 - 23.8	40	2.38	4.01	5.02
F42	8324	Kingswood	182	08.01 - 06.02	30	1.71	2.26	2.68
F42	2332	Hillfields	705	29.5 - 29.6	32	1.94	2.68	3.32
F46	3101	Knowle	853	18.6 - 19.7	32	3.38	4.06	4.75
F50	8325	Horfield	111	17.11 - 26.12	40	1.39	1.71	1.8
F50	8325	Horfield2	111	27.12 - 10.2	46	2.16	3.52	4.62
			12004			789		
Property Weighted						2.33	3.25	3.69
Acorn Weighted Totals							3.44	

Table 3 - Results of Legitimate Domestic Night Use Analysis